A Note on the Central Limit Theorem for the Eigenvalue Counting Function of Wigner Matrices

Sandrine Dallaporta, Van Vu

A NOTE ON THE CENTRAL LIMIT THEOREM FOR THE EIGENVALUE COUNTING FUNCTION OF WIGNER MATRICES

SANDRINE DALLAPORTA AND VAN VU

Abstract. The purpose of this note is to establish a Central Limit Theorem for the number of eigenvalues of a Wigner matrix in an interval. The proof relies on the correct aymptotics of the variance of the eigenvalue counting function of GUE matrices due to Gustavsson, and its extension to large families of Wigner matrices by means of the Tao and Vu Four Moment Theorem and recent localization results by Erdös, Yau and Yin.

1. Introduction

This note is concerned with the asymptotic behavior of the eigenvalue counting function, that is the number N_I of eigenvalues falling in an interval I, of families of Wigner matrices, when the size of the matrix goes to infinity. Wigner matrices are random Hermitian matrices M_n of size n such that, for $i < j$, the real and imaginary parts of $(M_n)_{ij}$ are iid, with mean 0 and variance $\frac{1}{2}$, $(M_n)_{ii}$ are iid with mean 0 and variance 1. An important example of Wigner matrices is the case where the entries are Gaussian, giving rise to the so-called Gaussian Unitary Ensemble (GUE). GUE matrices will be denoted by M_n'. In this case, the joint law of the eigenvalues is known, allowing for complete descriptions of their limiting behavior both in the global and local regimes (cf. for example [1]).

Denote by $\lambda_1, \ldots, \lambda_n$ the real eigenvalues of the normalized Wigner matrix $W_n = \frac{1}{\sqrt{n}} M_n$. The classical Wigner theorem states that the empirical distribution $\frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j}$ on the eigenvalues of W_n converges weakly almost surely as $n \to \infty$ to the semicircle law $d\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)dx$. Consequently, for any interval $I \subset \mathbb{R}$,

$$\frac{1}{n} N_I(W_n) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}_{\{\lambda_j \in I\}} \underset{n\to\infty}{\to} \rho_{sc}(I) \quad \text{almost surely.}$$

At the fluctuation level, it is known, due to the particular determinantal structure of the GUE, that

Theorem 1 (Costin-Lebowitz [2], Soshnikov [7] (see [1]))**.** *Let M_n' be a GUE matrix. Set $W_n' = \frac{1}{\sqrt{n}} M_n'$. Let I_n be an interval in $\mathbb{R}$. If* $\mathbf{Var}(N_{I_n}(W_n')) \underset{n\to\infty}{\to} \infty$,

V. Vu is supported by research grants from AFORS and NSF.

then

$$\frac{N_{I_n}(W_n') - \mathbf{E}[N_{I_n}(W_n')]}{\sqrt{\mathbf{Var}(N_{I_n}(W_n'))}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{1}$$

in distribution.

In 2005, Gustavsson [5] was able to fully describe for GUE matrices the asymptotic behavior of the variance of the counting function $N_I(W_n')$ for intervals $I = [y, +\infty)$ with $y \in (-2, 2)$ strictly in the bulk of the semicircle law. He established that

$$\mathbf{E}[N_I(W_n')] = n\rho_{sc}(I) + O\Big(\frac{\log n}{n}\Big) \quad \text{and} \quad \mathbf{Var}(N_I(W_n')) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n. \tag{2}$$

In particular therefore, if $I = [y, +\infty)$ with $y \in (-2, 2)$,

$$\frac{N_I(W_n') - \mathbf{E}[N_I(W_n')]}{\sqrt{\mathbf{Var}(N_I(W_n'))}} \underset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{3}$$

as well as

$$\frac{N_I(W_n') - n\rho_{sc}(I)}{\sqrt{\frac{1}{2\pi^2}\log n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{4}$$

(which we call below the CLT with numerics).

The purpose of this note is to extend these conclusions to non-Gaussian Wigner matrices. The class of Wigner matrices covered by our results is described by the following condition. Say that M_n satisfies condition ($\mathbb{C}0$) if the real part ξ and the imaginary part $\tilde{\xi}$ of $(M_n)_{ij}$ are independent and have an exponential decay: there are two constants C and C' such that

$$\mathbf{P}(|\xi| \geq t^C) \leq e^{-t} \quad \text{and} \quad \mathbf{P}(|\tilde{\xi}| \geq t^C) \leq e^{-t},$$

for all $t \geq C'$.

Say that two complex random variables ξ and ξ' match to order k if

$$\mathbf{E}\left[\mathrm{Re}(\xi)^m \mathrm{Im}(\xi)^l\right] = \mathbf{E}\left[\mathrm{Re}(\xi')^m \mathrm{Im}(\xi')^l\right]$$

for all $m, l \geq 0$ such that $m + l \leq k$.

The following theorem is the main result of this note.

Theorem 2. *Let M_n be a random Hermitian matrix whose entries satisfy condition ($\mathbb{C}0$) and match the corresponding entries of GUE up to order 4. Set $W_n = \frac{1}{\sqrt{n}}M_n$. Then, for any $y \in (-2, 2)$ and $I(y) = [y, +\infty)$, setting $Y_n := N_{I(y)}(W_n)$, we have*

$$\mathbf{E}[Y_n] = n\rho_{sc}(I(y)) + o(1) \quad \textit{and} \quad \mathbf{Var}(Y_n) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n,$$

and the sequence (Y_n) satisfies the CLT

$$\frac{Y_n - \mathbf{E}[Y_n]}{\sqrt{\mathbf{Var}(Y_n)}} \underset{n\to\infty}{\to} \mathcal{N}(0,1).$$

The theorem is established in the next two sections. In a first step, relying on Gustavsson's results and its extension to Wigner matrices by Tao and Vu [9], we establish that (Y_n) satisfies the CLT with numerics (4). In a second step, we use recent results of Erdös, Yau and Yin [3] on the localization of eigenvalues in order to prove that $\mathbf{E}[Y_n]$ and $\mathbf{Var}(Y_n)$ are close to those of M'_n (GUE) and therefore satisfy (2).

2. CLT with numerics and eigenvalues in the bulk

On the basis of the CLT with numerics, Gustavsson [5] described the Gaussian behavior of eigenvalues in the bulk of the semicircle law in the form of

$$\sqrt{\frac{4-t(i/n)^2}{2}}\,\frac{\lambda_i(W'_n)-t(i/n)}{\frac{\sqrt{\log n}}{n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{5}$$

in distribution, where $t(x)\in[-2,2]$ is defined for $x\in[0,1]$ by

$$x=\int_{-2}^{t(x)} d\rho_{sc}(t)=\frac{1}{2\pi}\int_{-2}^{t(x)}\sqrt{4-x^2}dx.$$

More informally, $\lambda_i(W'_n)\approx t(i/n)+\mathcal{N}(0,\frac{2\log n}{(4-t(i/n)^2)n^2})$. This is achieved by the tight relation between eigenvalues and the counting function expressed by the elementary equivalence, for $I(y)=[y,+\infty)$, $y\in\mathbb{R}$,

$$N_{I(y)}(W_n)\le n-i \quad \text{if and only if} \quad \lambda_i\le y. \tag{6}$$

The result (5) was extended in [9] to large families of Wigner matrices satisfying condition ($\mathbb{C}$0) by means of the Four Moment Theorem (see [9] and [10]). Now using the reverse strategy based on (6), (5) may be shown to imply back the CLT with numerics (4) for Wigner matrices whose entries match those of the GUE up to order 4. We provide some details in this regard relying on the following simple consequence of the Four Moment Theorem.

Proposition 3. *Let M_n and M''_n be two random matrices satisfying condition* ($\mathbb{C}$0) *such that their entries match up to order* 4. *There exists $c>0$ such that, if n is large enough, for any $y\in(-2,2)$ and any $(i,j)\in\{1,\dots,n\}^2$, if $I(y)=[y,+\infty)$,*

$$\big|\mathbf{P}(\lambda_i\in I(y))-\mathbf{P}(\lambda''_i\in I(y))\big|\le n^{-c},$$

and

$$\big|\mathbf{P}(\lambda_i\in I(y)\wedge\lambda_j\in I(y))-\mathbf{P}(\lambda''_i\in I(y)\wedge\lambda''_j\in I(y))\big|\le n^{-c}.$$

As announced, we would like to show that the behavior of eigenvalues in the bulk (5) extended to Wigner matrices leads to the CLT with numerics for such matrices, namely,

$$\frac{N_{I(y)}(W_n)-n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1), \tag{7}$$

in distribution for Wigner matrices W_n satisfying ($\mathbb{C}0$). To prove this, observe that for every $x \in \mathbb{R}$.

$$\mathbf{P}\Big(\frac{N_{I(y)}(W_n) - n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \leq x\Big) = \mathbf{P}(N_{I(y)}(W_n) \leq n - i_n)$$

where $i_n = n\rho_{sc}((-\infty, y]) - x\sqrt{\frac{1}{2\pi^2}\log n}$. Then, by (6),

$$\begin{aligned}\mathbf{P}\Big(\frac{N_{I(y)}(W_n) - n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \leq x\Big) &= \mathbf{P}(\lambda_{i_n}(W_n) \leq y) \\ &= \mathbf{P}\Big(\sqrt{\frac{4 - t(i_n/n)^2}{2}}\frac{\lambda_{i_n}(W_n) - t(i_n/n)}{\frac{\sqrt{\log n}}{n}} \leq x_n\Big),\end{aligned}$$

where $x_n = \sqrt{\frac{4-t(i_n/n)^2}{2}}\frac{y-t(i_n/n)}{\frac{\sqrt{\log n}}{n}}$. Now $\frac{i_n}{n} \to \rho_{sc}((-\infty, y]) \in (0, 1)$. Furthermore, $x_n \to x$ since

$$\begin{aligned} t(i_n/n) &= t\Big(\rho_{sc}((-\infty, y]) - \frac{x}{n}\sqrt{\frac{1}{2\pi^2}\log n}\Big) \\ &= t\Big(\rho_{sc}((-\infty, y])\Big) - t'\Big(\rho_{sc}((-\infty, y])\Big)\frac{x}{n}\sqrt{\frac{1}{2\pi^2}\log n} + o\Big(\frac{\sqrt{\log n}}{n}\Big) \\ &= y - x\sqrt{\frac{2}{4-y^2}}\frac{\sqrt{\log n}}{n} + o\Big(\frac{\sqrt{\log n}}{n}\Big). \end{aligned}$$

Hence $\frac{y-t(i_n/n)}{\frac{\sqrt{\log n}}{n}} = x\sqrt{\frac{2}{4-y^2}} + o(1)$, from which $x_n \to x$.

Applying (5) (extended to Wigner matrices), we obtain that

$$\mathbf{P}\Big(\sqrt{\frac{4 - t(i_n/n)^2}{2}}\frac{\lambda_{i_n}(W_n) - t(i_n/n)}{\frac{\sqrt{\log n}}{n}} \leq x_n\Big) \underset{n\to\infty}{\to} \mathbf{P}(X \leq x),$$

where $X \sim \mathcal{N}(0, 1)$, implying (7).

3. Estimating the mean and the variance of Y_n

To reach the CLT of Theorem 2 from the CLT with numerics (7), it is necessary to suitably control the expectation and variance $\mathbf{E}[Y_n]$ and $\mathbf{Var}(Y_n)$ of the eigenvalue counting function, and to show that their behaviors are identical to the ones for GUE matrices. The direct use of the Four Moment Theorem is unfortunately not enough to this purpose since it only provides proximity on a small number of eigenvalues. But recent results of Erdös, Yau and Yin [3], presented in the following statement, describe strong localization of the eigenvalues of Wigner matrices which provides the additional step necessary to complete the argument.

Theorem 4. *Let M_n be a random Hermitian matrix whose entries satisfy condition* ($\mathbb{C}0$). *There is a constant $C > 0$ such that, for any $i \in \{1, \ldots, n\}$,*

$$\mathbf{P}(|\lambda_i - t(i/n)| \geq (\log n)^{C\log\log n}\min(i, n-i+1)^{-1/3}n^{-2/3}) \leq n^{-3}.$$

Note that if $n\varepsilon \leq i \leq (1-\varepsilon)n$ for some small $\varepsilon > 0$, then $\min(i, n-i+1) \geq n\varepsilon$ so that

$$\mathbf{P}\Big(|\lambda_i - t(i/n)| \geq n^{-1}\varepsilon^{-1/3}(\log n)^{C\log\log n}\Big) \leq n^{-3}. \tag{8}$$

The next lemma presents the main conclusion on the expectation and variance of the eigenvalue counting function, extending Gustavsson's conclusion (2) for the GUE to Wigner matrices of the class ($\mathbb{C}0$). Once this lemma is established, Theorem 2 will follow.

Lemma 5. *Set $W_n = \frac{1}{\sqrt{n}}M_n$, $I(y) = [y, +\infty)$ where $y \in (-2, 2)$, and $Y_n = N_{I(y)}(W_n)$. Then*

$$\mathbf{E}[Y_n] = n\rho_{sc}(I(y)) + o(1) \quad \textit{and} \quad \mathbf{Var}(Y_n) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n.$$

Proof. By Gustavsson's results (2) therefore, if Y_n' denotes $N_{I(y)}(W_n')$ in the case M_n' is GUE,

$$\mathbf{E}[Y_n'] = n\rho_{sc}(I(y)) + O\Big(\frac{\log n}{n}\Big) \quad \text{and} \quad \mathbf{Var}(Y_n') = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n.$$

Hence, to establish Lemma 5, it suffices to show that $\mathbf{E}[Y_n] = E[Y_n'] + o(1)$ and $\mathbf{Var}(Y_n) = \mathbf{Var}(Y_n') + o(1)$. Below, we only deal with the variance, the argument for the expectation being similar and actually simpler.

Set $A_i = \mathbf{1}_{\{\lambda_i \in I\}}$, for $i \in \{1, \dots, n\}$. Notice that

$$|\mathbf{Var}(Y_n) - \mathbf{Var}(Y_n')| \leq \sum_{1\leq i,j\leq n} |(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]) - (\mathbf{E}[A_i'A_j'] - \mathbf{E}[A_i']\mathbf{E}[A_j'])|.$$

Call an index i *first class* if $\mathbf{E}[A_i] \geq 1 - n^{-3}$ or $\leq n^{-3}$ and *second class* otherwise.

Note that if j is first class, then, for all $i \in \{1, \dots, n\}$,

$$|\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]| = O(n^{-3}).$$

Indeed, if $\mathbf{E}[A_j] \leq n^{-3}$, then both terms between the absolute value signs are $O(n^{-3})$, so that $|\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]| = O(n^{-3})$. The other case can be brought back to this case by the identity

$$|\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]| = |\mathbf{E}[B_iB_j] - \mathbf{E}[B_i]\mathbf{E}[B_j]|,$$

where B_i is the complement of A_i. Consequently,

$$\sum_{\substack{i \text{ or } j \\ 1^{\text{st}} \text{ class}}} |\big(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]\big) - \big(\mathbf{E}[A_i'A_j'] - \mathbf{E}[A_i']\mathbf{E}[A_j']\big)| = O(n^{-1}). \tag{9}$$

Theorem 4 shows that there are only $O((\log n)^{C\log\log n})$ second class indices. Indeed, set $\eta_n = n^{-1}\varepsilon^{-1/3}(\log n)^{C\log\log n}$ and suppose first that $i \in \{1, \dots, n\}$ is such that $t(i/n) < y - \eta_n$:

- if $t(i/n) > t(\varepsilon)$, (8) is true for W_n. Then
$$\mathbf{P}(\lambda_i \in I_n) \leq \mathbf{P}\Big(|\lambda_i - t(i/n)| \geq \eta_n\Big) \leq n^{-3}.$$
- if $t(i/n) < t(\varepsilon)$, choose j such that $t(\varepsilon) < t(j/n) < y - \eta_n$ (take ε small enough and n large enough such that there is such a j). Then $\lambda_i \leq \lambda_j$ and $\mathbf{P}(\lambda_i \in I_n) = \mathbf{P}(\lambda_i \geq y) \leq \mathbf{P}(\lambda_j \geq y) = \mathbf{P}(\lambda_j \in I_n) \leq n^{-3}$.

Similarly one can show that if $i \in \{1, \ldots, n\}$ is such that $t(i/n) > y + \eta_n$, then i is first class.

As a consequence of this discussion, $i \in \{1, \ldots, n\}$ can only be second class if $y - \eta_n < t(i/n) < y + \eta_n$. We need to count these possible i's. By definition of $t(i/n)$, $i = \frac{n}{2\pi}\int_{-2}^{t(i/n)} \sqrt{4-x^2}dx$. Thus,
$$\frac{n}{2\pi}\int_{-2}^{y-\eta_n} \sqrt{4-x^2}dx \leq i \leq \frac{n}{2\pi}\int_{-2}^{y+\eta_n} \sqrt{4-x^2}dx.$$
In this case i belongs to an interval of length
$$\frac{n}{2\pi}\int_{y-\eta_n}^{y+\eta_n} \sqrt{4-x^2}dx \leq \frac{2}{\pi\varepsilon^{1/3}}(\log n)^{C\log\log n}.$$

Therefore, there are at most $\frac{2}{\pi\varepsilon^{1/3}}(\log n)^{C\log\log n} + 1 = O((\log n)^{C\log\log n})$ second class i's.

Next, by Proposition 3, it is easily seen that if both i, j are second class, then
$$|(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]) - (\mathbf{E}[A'_iA'_j] - \mathbf{E}[A'_i]\mathbf{E}[A'_j])| = O(n^{-c})$$
for some positive constant c. Since the number of such pairs is $O((\log n)^{2C\log\log n})$, we have

(10)
$$\sum_{\substack{i \text{ and } j \\ 2^{\text{nd}} \text{ class}}} |\big(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]\big) - \big(\mathbf{E}[A'_iA'_j] - \mathbf{E}[A'_i]\mathbf{E}[A'_j]\big)| = O(n^{-c}(\log n)^{2C\log\log n}).$$

To conclude,
$$\begin{aligned}|\mathbf{Var}(Y_n) - \mathbf{Var}(Y'_n)| \leq &\sum_{\substack{i \text{ or } j \\ 1^{\text{st}} \text{ class}}} |\big(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]\big) - \big(\mathbf{E}[A'_iA'_j] - \mathbf{E}[A'_i]\mathbf{E}[A'_j]\big)| \\ &+ \sum_{\substack{i \text{ and } j \\ 2^{\text{nd}} \text{ class}}} |\big(\mathbf{E}[A_iA_j] - \mathbf{E}[A_i]\mathbf{E}[A_j]\big) - \big(\mathbf{E}[A'_iA'_j] - \mathbf{E}[A'_i]\mathbf{E}[A'_j]\big)|,\end{aligned}$$
so that (9) and (10) lead to
$$|\mathbf{Var}(Y_n) - \mathbf{Var}(Y'_n)| \leq O(n^{-1}) + O(n^{-c}(\log n)^{2C\log\log n}) = o(1),$$
as claimed. This shows that $\mathbf{Var}(Y_n) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n$. As mentioned earlier, it may be shown similarly that $\mathbf{E}[Y_n] = n\rho_{sc}(I(y)) + o(1)$ and the proof of Lemma 5 is thus complete. □

4. About real Wigner matrices

In this section, we briefly indicate how the preceding results for Hermitian random matrices may be stated similarly for real Wigner symmetric matrices. To this task, we follow the same scheme of proof, relying in particular on the corollary of Tao and Vu Four Moment Theorem (Proposition 3) which also holds in the real case (cf. [6]).

Real Wigner matrices are random symmetric matrices M_n of size n such that, for $i < j$, $(M_n)_{ij}$ are iid, with mean 0 and variance 1, $(M_n)_{ii}$ are iid with mean 0 and variance 2. As in the complex case, an important example of real Wigner matrices is the case where the entries are Gaussian, giving rise to the so-called Gaussian Orthogonal Ensemble (GOE).

The main issue is actually to establish first the conclusions for the GOE. This has been suitably developed by O'Rourke in [6] by means of interlacing formulas (cf. [4]).

Theorem 6 (Forrester-Rains)**.** *The following relation holds between matrix ensembles:*

$$\mathrm{GUE}_n = \mathrm{even}(\mathrm{GOE}_n \cup \mathrm{GOE}_{n+1}).$$

This statement can be interpreted in the following way. Take two independent matrices from the GOE, one of size n and the other of size $n+1$. If we surimperpose the $2n+1$ eigenvalues on the real line and then take the n even ones, they have the same distribution as the eigenvalues of a $n \times n$ matrix from the GUE.

Let I be an interval in $\mathbb{R}$. Let $M_n^{\mathbb{R}}$ be a GOE matrix and $M_n^{\mathbb{C}}$ be a GUE matrix. $W_n^{\mathbb{R}}$ and $W_n^{\mathbb{C}}$ are the corresponding normalized matrices. The preceding interlacing formula leads to

- $\mathbf{E}[N_I(W_n^{\mathbb{R}})] = \mathbf{E}[N_I(W_n^{\mathbb{C}})] + O(1)$
- $\mathbf{Var}(N_I(W_n^{\mathbb{R}})) = 2\mathbf{Var}(N_I(W_n^{\mathbb{C}})) + O(1)$, if $\mathbf{Var}(N_I(W_n^{\mathbb{C}})) \underset{n\to\infty}{\to} \infty$.

Relying on this result and on the GUE case, O'Rourke proved the following theorem:

Theorem 7. *Let $M_n^{\mathbb{R}}$ be a GOE matrix. Set $W_n^{\mathbb{R}} = \frac{1}{\sqrt{n}} M_n^{\mathbb{R}}$. Then, for any $y \in (-2,2)$ and $I(y) = [y, +\infty)$, setting $Y_n^{\mathbb{R}} := N_{I(y)}(W_n^{\mathbb{R}})$, we have*

$$\mathbf{E}[Y_n^{\mathbb{R}}] = n\rho_{sc}(I(y)) + O(1) \quad \textit{and} \quad \mathbf{Var}(Y_n^{\mathbb{R}}) = \Big(\frac{1}{\pi^2} + o(1)\Big)\log n,$$

and the sequence $(Y_n^{\mathbb{R}})$ satisfies the CLT

$$\frac{Y_n^{\mathbb{R}} - \mathbf{E}[Y_n^{\mathbb{R}}]}{\sqrt{\mathbf{Var}(Y_n^{\mathbb{R}})}} \underset{n\to\infty}{\to} \mathcal{N}(0,1).$$

Following exactly the same scheme as for complex Wigner matrices leads to the same conclusion: Theorem 7 is true for Wigner symmetric matrices, provided their entries match the corresponding entries of GOE up to order 4 and satisfy condition ($\mathbb{C}0$).

The CLT for the eigenvalue counting function has been investigated as well for families of covariance matrices. The main conclusion of this work holds similarly in this case conditioned however on the validity of the Erdös-Yau-Yin rigidity theorem for covariance matrices. There is strong indication that the current approach by Erdös, Yin and Yau for Wigner matrices will indeed yield such a result. All the other ingredients of the proof are besides available. Indeed, Su (cf. [8]) carried out computations for Gaussian covariance matrices and proved both the CLT and the correct asymptotics for mean and variance. Tao and Vu in [11] extended their Four Moment Theorem to such matrices. Arguing then as for Wigner matrices, we could reach in the same way a CLT with numerics for suitable families of covariance matrices.

References

[1] G. Anderson, A. Guionnet, O. Zeitouni, An Introduction to Random Matrices, 2010.
[2] O. Costin, J. L. Lebowitz, Gaussian Fluctuations in Random Matrices, Phys. Rev. Lett. 75 (1995), p. 69-72.
[3] L. Erdös, H-T. Yau, J. Yin, Rigidity of Eigenvalues of Generalized Wigner Matrices, arXiv:1007.4652, 2010.
[4] P. Forrester, E. Rains, Inter-Relationships between Orthogonal, Unitary and Symplectic Matrix Ensembles, Cambridge University Press, Cambridge, United Kingdom (2001), p. 171-208.
[5] J. Gustavsson, Gaussian Fluctuations of Eigenvalues in the GUE, Ann. I. Poincaré - PR 41 (2005), p. 151-178.
[6] S. O'Rourke, Gaussian Fluctuations of Eigenvalues in Wigner Random Matrices, 2009.
[7] A. Soshnikov, Gaussian Fluctuation for the Number of Particles in Airy, Bessel, Sine, and other Determinantal Random Point Fields, J. Statist. Phys. 100 (2000), 3-4, p. 491-522.
[8] Z. Su, Gaussian Fluctuations in Complex Sample Covariance Matrices, Electronic Journal of Probability 11 (2006), p. 1284-1320.
[9] T. Tao and V. Vu, Random Matrices: Universality of Local Eigenvalues Statistics, to appear in Acta Math, arXiv:0906.0510.
[10] T. Tao and V. Vu, Random Matrices: Universality of Local Eigenvalue Statistics up to the Edge, Comm. Math. Phys. 298 (2010), 2, p. 549-572.
[11] T. Tao and V. Vu, Random Covariance Matrices: Universality of Local Statistics of Eigenvalues, arXiv:0912.0966, 2010.

Institut de Mathématiques de Toulouse, UMR 5219 du CNRS, Université de Toulouse, F-31062 Toulouse, France

E-mail address: sandrine.dallaporta@math.univ-toulouse.fr

Department of Mathematics, Rutgers, Piscataway, NJ 08854

E-mail address: vanvu@math.rutgers.edu